Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol.

Introduction

This appendix includes a compilation of memos and analysis completed by Kathi Irvine, PhD, in the Department of Math Science from Montana State University. The Klamath Network (KLMN) worked with Kathi to develop R code that could be used to determine the number of routes and years of sampling needed to achieve 80% power to detect a 50% decline in relative abundance for an individual species after 20 years based on the monitoring strategy described in the KLMN Landbird Monitoring Protocol.

The first memo dated October 13, 2009 (see below) provides a description of the models used as part of our power analysis. In addition, this memo provides instructions on how to use the R Code and a few examples of the output from the code.

The second memo dated October 28, 2009 (see below) provides the results of two examples (Spotted Towhee at LABE and Pacific-Slope Flycatcher at RNSP) of our ability to have enough power to detect trends as described in the protocol narrative. As stated in the memo, the site-by-year variance (sig2c) plays a significant role in determining the results of the models. If sig2c is estimated to be 0.05, with 25 sites at Lava Beds National Monument (LABE), we would have 91.4% power to detect a 50% decline after 20 years within 10 years (three sampling events), which exceeds our goal of 80% power. However, if we estimate sig2c as 1.00, then after 10 years we would have 40.5% power to detect a 50% decline after 20 years. It would take 30 years (10 sampling events) to reach our goal of 80% power.

At this point in time, the KLMN has completed one sampling event at two (LABE, RNSP) of the six parks where we plan on implementing this project. Since we have only completed one visit, we had to bring in additional data to try and estimate the site-by-year variance described above. In an effort to make the best determination of the site-by-year variance, we examined the Breeding Bird Survey data from four routes that were near LABE. This effort is described in the third memo dated November 3, 2009, which is provided below. As the memo states, in some cases, species that have a high site-by-year variability make it difficult to detect trends. This is shown in example 1 (Spotted Towhee) and 2 (Brown Headed Cowbird), where it is predicted it will take 30+ years (10+ sampling events) to be able to detect trends in these species. However, for species that have very little site-by-year variance, we will be able to detect trends in as little as 15 years (five sampling events), as in shown in example 3 for the Western Meadowlark.

Assumptions

It is assumed that the BBS routes that surround our parks cover a larger area (routes are 24.5 miles long) and bisect more heterogeneous habitats than the routes we have established within the park. Based on the design of the route layout in this protocol, we are assuming the site-by-year variance will be less variable than the BBS routes. If this holds true, the power estimates in this analysis are underestimates. In order to test this assumption, a power analysis will be completed as part of our third Analysis and Synthesis report. Although we have already shown we have the power to detect trends in common species that have little site-by-year variance

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

within 15 years, this analysis will provide us with a list of species that we can potentially provide trend data on as part of the fifth Analysis and Synthesis report.

In addition, as described in SOP 17: Data Analysis and Reporting, in the first Analysis and Synthesis report, we will be developing detection functions for individual species. A detection function was not used for the analysis described in the memos below. Instead, any species observation that was beyond 50 meters from the listening station was dropped. It is likely that developing detection functions will help provide better estimates of relative abundance, which will improve our power to detect trends. Since detection functions were not created, there is the chance the power estimates provided in these examples are underestimates. Similar to the above assumption, we will test this assumption when completing the power analysis in the third Analysis and Synthesis report.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

MEMO 1: October 13, 2009

October 13, 2009
Prepared by K. Irvine, PhD

Dept of Math Sciences
Montana State University
Bozeman, MT 59715
irvine@math.montana.edu
(406) 994-5357

Consultation with the author is recommended before using the code.

This document pertains to the following R code files: powercommands051808, Rcode_CalcPowerBIRDS, Rcode_CalcPowerRandom, Rcode_SimPopnChangeBIRDS, Rcode_SimPopnChangeRandomBIRDS, and Satterthwaite_DF_functions. Specific documentation is provided within the code itself. These files were used to produce the May 18, 2008 memo titled "Trend Power Analysis for Landbirds Monitoring Protocol for Klamath Network Parks."

The file powercommands051808 contains all the R commands that were used for the memo results. The other files (essentially R functions) must be read into R in order for the commands to work properly.

The code is specific to Model 1 and Model 2 described in the memo. If a different model is used to estimate the variance components from pilot data, then the power code is essentially worthless. Also, the degrees of freedom used for the test of linear trend is an approximation to the Satterthwaite method as described in the Giesbrecht and Burns (1985). This approximation could improve in time as statistical research progresses. Further the trend model is dependent on the species of interest. Thus a decision needs to be made based on pilot data as to the particular variance component structure. In other words, the power code should not be used blindly once the NPS land birds protocol has been implemented for several years and it is time to perform the power analysis and submit report 3 as outlined in SOP 17.

F. G. Giesbrecht and J. C. Burns 1985 **Two-Stage Analysis Based on a Mixed Model: Large-Sample Asymptotic Theory and Small-Sample Simulation Results** *Biometrics*, Vol. 41, No. 2 pp. 477-486

Instructions for Use of Code

Option 1: Random effect of sites

If you want to investigate the number of routes effect on power, you will need to assume a random effect of site. The following files are needed:

1. Rcode_CalcPowerRandom.txt
2. Rcode_SimPopnChangeRandomBIRDS.txt
3. Satterthwaite_DF_functions.txt

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

Option 2: Fixed effect of sites

This option is when you already have pilot data for the sites where you plan on sampling. OR the number of routes is fixed and you are only interested in the effect of number of years of sampling on power. You need the following files:

1. Rcode_CalcPowerBIRDS
2. Rcode_SimPopnChangeBIRDS
3. Satterthwaite_DF_functions

Instructions for either option:

Step 1: Change the directory that R is using as the default. File menu → Change dir…-→ select the folder where your files are located

Step 2: The main file that contains the commands used in the memo is powercommands051808.txt. Open this file within the R package by using File→ Open Script → select "powercommands051808.txt"

Step 3: To submit all the functions needed to estimate the power; open all the files within the R program by following step 2 for the following files listed under option 1 or 2.

If using Option 1:

Step 4: To estimate the power for either 25, 30, or 35 number of routes and different number of visits (=sampling periods). You will need the following things as inputs to the function CalcPowerRandom

alpha,its,nosites,yrs,StartYear,change,sig2a,sig2b, sig2t,and sig2c defined as follows:

alpha desired type I error rate
its number of iterations for simulated power
nosites is the number of sites(=routes)
yrs number of years of sampling: revisit design so if every 3 years 0,3,6 etc
** change is the percent decline over 20 years of interest : 50% decline would be -.50
sig2a is the variance component for random intercept
sig2b is the year to year variability
sig2t is the variance component for the random slopes
sig2c is the error variance of site*year variance component

**The code is currently set-up to only investigate only a percent change over 20 years, so say 50% decline or 30% decline over 20 years. This can be changed such that you can change the interval over which you want the decline/increase to occur.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

EXAMPLE

#Power estimated based on 50% decline over 20 years for 40 routes, alpha=.10, and different sampling durations

CalcPowerRandom(.10,1000,40,c(0,3,6,9,12,15,18,21,24,27),2004,-.5,sig2a=.023,sig2b=.21, sig2t=.003,sig2c=.047)

#Power estimated based on 50% decline over 38 years for 40 routes, alpha=.10, and different sampling durations

CalcPowerRandom(.10,1000,40,c(seq(0,38,by=2)),2004,-.5,sig2a=.023,sig2b=.21, sig2t=.003,sig2c=.047)

The estimates from the variance components are based on pilot data. In order to estimate sig2a, sig2b, sig2t, and sig2c you will need at least two years of data. However, for a start I don't think there is harm in using the estimates of sig2c from the KBO data for a species and the first year of data for sig2a to estimate the site-to-site variability within a given park.

If using Option 2:

Step 4: You will need the following things as inputs to the CalcPowerBIRDS function:

alpha,its,nosites, yrs, StartYear, change, ai,sig2b, sig2t,sig2c

defined as follows:

 alpha desired type I error rate;
 its number of iterations for simulated power;
 nosites is the number of sites(=routes);
 yrs number of years of sampling: revisit design so if every 3 years 0,3,6 etc ;
 **change is the percent decline over 20 years of interest : 50% decline would be -.50;
 ai are the fixed site intercepts;
 sig2b is the year to year variability;
 sig2t is the variance component for the random slopes;
 sig2c is the error variance of site*year variance component.

EXAMPLE:

#"fixsite" are fixed site intercepts estimated from the pilot dataset in SAS for model 1

fixsite<-c(0.5311834, 0.8897791,0.5603643,0.8182710,0.3307078,0.5493261,0.7564269,0.9434209,0.8915809, 0.9087020, 0.4266378)

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

#Power estimated based on a 50% decline over 20 years with 4 sampling periods (every 3 years) for 11 routes.

CalcPowerBIRDS(.10,1000,11,c(0,3,6,9),2004,-.5, fixsite,sig2b=.008, sig2t=.004,sig2c=.0467)

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

MEMO 2: October 28, 2009

October 28, 2009
Prepared by K. Irvine, PhD

Dept of Math Sciences
Montana State University
Bozeman, MT 59715
irvine@math.montana.edu
(406) 994-5357

In this document, I investigate the number of routes and years of sampling to achieve 80% power to detect a 50% decline in relative abundance of Spotted Towhee after 20 years in Lava Beds National Monument. Also, the second example is for the Pacific-slope Flycatcher in Redwood National Park.

I assume that the routes are fixed. If some sort of panel design was used, the code should be changed to reflect that revisit design.

The example used the pilot data provided by Sean Mogren: *PointCounts.xls.* The following used the R code developed previously and tested using KBO data.

Example 1: Spotted Towhee Lava Beds

The code for the Spotted Towhee examples is provided within the file: PilotDataCommands_SPTO.txt.

In order to run the wrapper function POWER.RandomInterceptsModel() in step 6, the following files need to be submitted to the R console in the following order
   1. Rcode_SimPopnChangeRandomBIRDS
   2. Satterthwaite_DF_functions
   3. Rcode_CalPowerRandom
   4. Rcode_PowerFunction

The steps followed in the example file PilotDataCommands_SPTO.txt.

   1) Create a .csv file to import into R (done outside of R in excel). Import csv file into R.
   2) Subset data to one park and one species of interest
   3) Calculate relative abundance for each species at a route as # birds detected <50m/ # points. One issue is that the denominator varies for Redwoods. To accommodate that, I have included additional code needed for the Redwoods sites at this step. See PilotDataCommands_PSFL.txt code for that example.
   4) Look at the data to assess assumption of normality. For instance, with SPTO the histogram on the log scale appears ok. This is hard to assess with minimal pilot data, but needs to be done before proceeding with a power analysis once more data is available.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

5) Calculate the site-to-site variance observed in the pilot data. Calculate the estimate of mu.
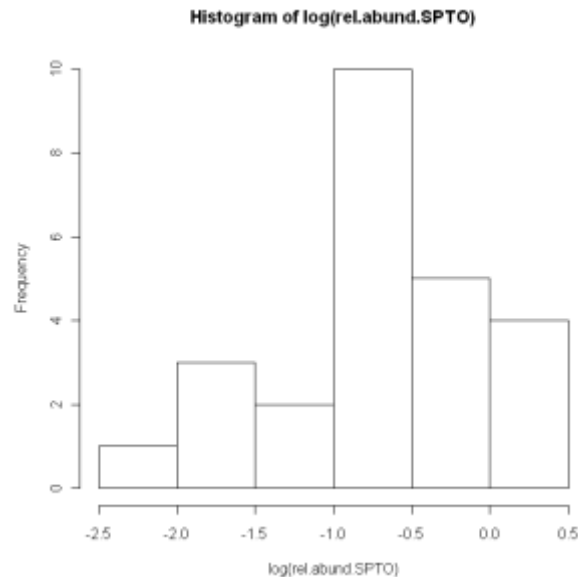6) Run the power function for one species and park based on a set of inputs.

Figure 1. Histogram of relative abundance on the log scale for the Spotted Towhee used to examine the assumption of normality.

For step 6, you need to decide on the alpha-level for the trend test, the number of iterations for the simulation, the range of sample sizes of interest (e.g., 20,25,30 sites), the number of years between sampling (off years), the range of the number of years of sampling of interest (5, 10, 20, 30 years of sampling), the start year (not that critical), the % decrease/increase in median relative abundance after 20 years, variance component estimates.

I used the following in the SPTO example:

```
#.10 = alpha level for trend test
# 500 is the number of iterations for each simulation
# c(20,25,30) is the number of sites
# 2 is the number of off years between sampling 0,3,6,9 for 10 years period of sampling
# c(5,10,20,30) is the number of years of sampling
# 2008 start year
# -50% decrease over 20 years
# mu is the estimated overall mean across sites
# sig2a is the variance component for random intercept
# sig2b is the year to year variability
# sig2t is the variance component for the random slopes
# sig2c is the error variance of site*year variance component
```

One issue I ran into is the estimate of the site*year variance component. A range of variance estimates should be considered. Also, because of the general model formulation in the code, I got

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

the following error message "In mer_finalize(ans) ... : singular convergence (7)" which is probably due to the zero variance component for year. The results should still yield valid estimates of power under the assumed model and given variance inputs.

Step 6: Example Results Output SPTO Lava Beds:

OUT<-POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-.64,sig2a=.491,sig2b=0, sig2t=0,sig2c=.05)

	Years			
Sites	**5**	**10**	**20**	**30**
20	0	**0.814**	0.972	0.980
25	0	0.914	0.992	0.996
30	0	0.956	0.994	1.000

The table is set-up such that reading across a row indicates the power for the given *row* sample size for different sampling durations (columns). For example, the power to detect a 50% decline after 20 years is 81.4% with 10 years of sampling (2008-2018) every 3rd year with 20 fixed routes; 97.2% with 20 years of sampling (2008-2028); 98% with 30 years of sampling (2008-2038). This is based on estimated variance components, and may be an overestimated if the variances are too small.
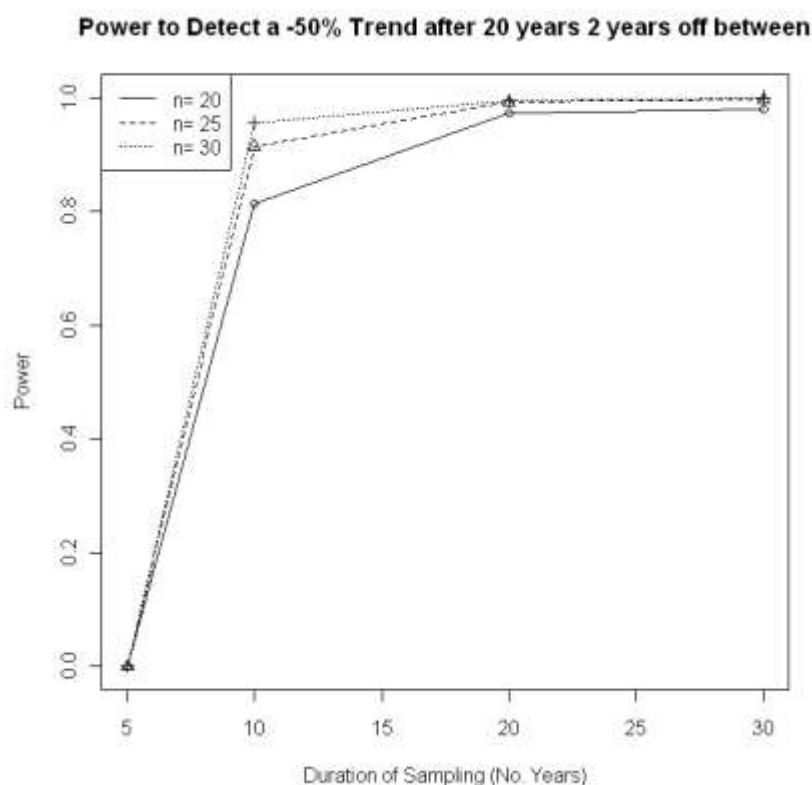

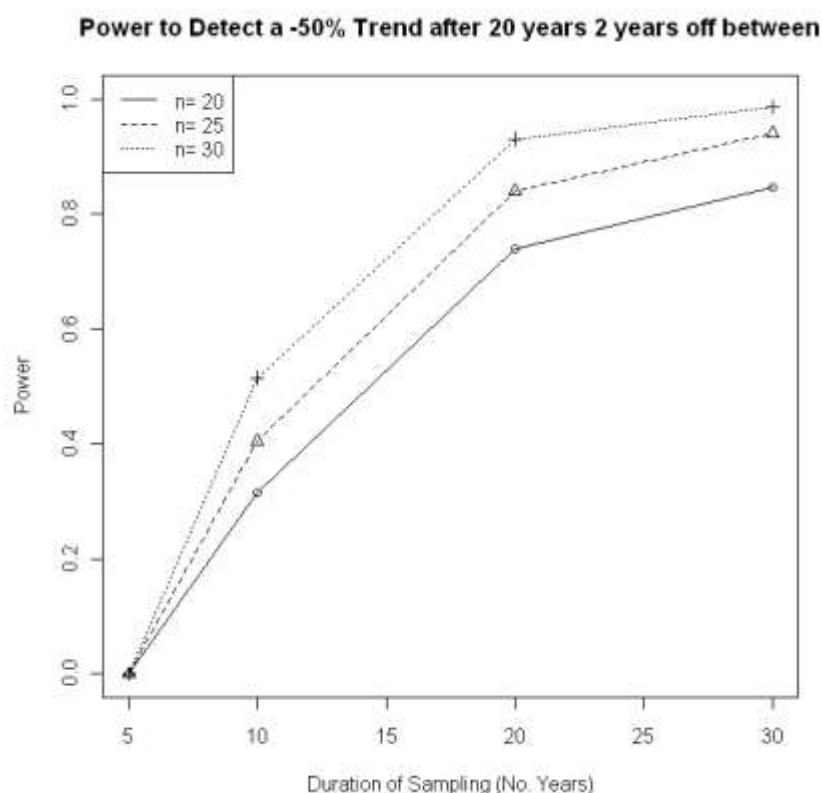**Power to Detect a -50% Trend after 20 years 2 years off between**

Figure 2. Power to detect trends in Spotted Towhee over a 30 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 0.05.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

For another example, to demonstrate the role of the estimated variance component for sig2c, I changed the input to 1 from .05, the power decreases as expected. The power to detect a 50% decline after 20 years is 31.5% with 10 years of sampling (2008-2018) every 3rd year with 20 fixed routes as opposed to 81.4%.

Sites	5	10	20	30
20	0	0.315	0.74	0.845
25	0	0.405	0.84	0.940
30	0	0.515	0.93	0.985

Figure 3. Power to detect trends in Spotted Towhee over a 30 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 1.00.

Example 2: Pacific-Slope Flycatcher Redwood National Park

The code for this example is in the file "PilotDataCommands_PSFL.txt."

Step 3:

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

Notice Step 3 for calculating relative abundance is the only difference in the code for this example compared to the SPTO example.

**When I calculate the number of points sampled for each site, I find there were 7 sampled at site RW12; however, in the annual report Table 3 says there were only 6. My calculations correspond with the "Route and Points.xls" file as well.
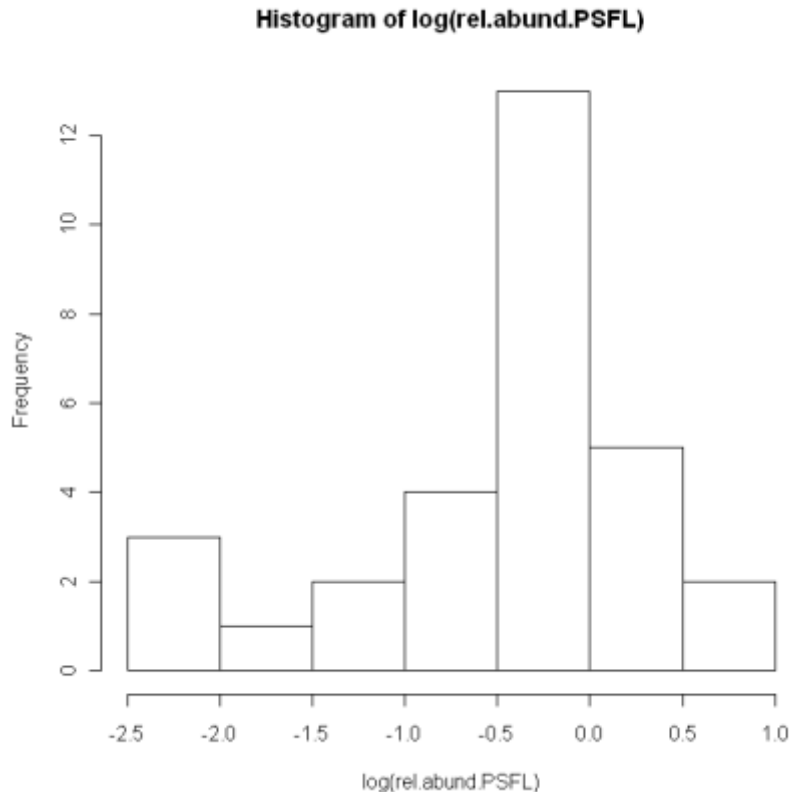
Step 4:

Figure 4. Histogram of relative abundance on the log scale for the Pacific-Slope Flycatcher used to examine the assumption of normality.

Step 5: Pilot Values calculated for PSFL in Redwoods

sig2a = .611; mu = -.45

Step 6: Example Results Output PSFL Redwoods:

```
OUT.PSFL1<-POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-.45,sig2a=.611,sig2b=0,
sig2t=0,sig2c=1)

5  10  20  30
20 0 0.132 0.402 0.598
25 0 0.124 0.544 0.630
```

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).
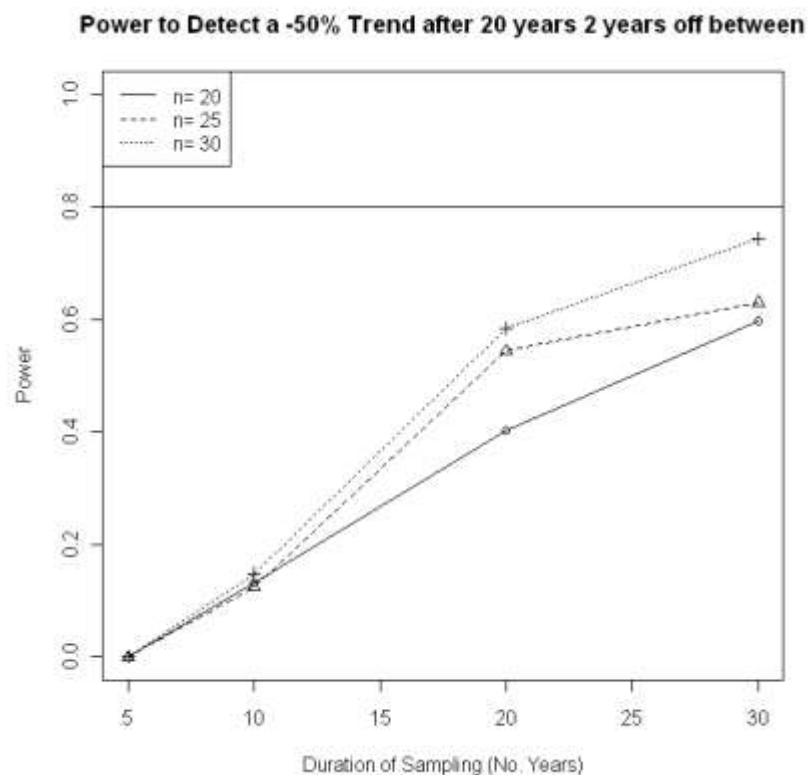
30 0 0.148 0.584 0.744

Figure 5. Power to detect trends in Pacific-Slope Flycatchers over a 30 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 0.05.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol.

Appendix 1.

R-Code for example SPTO

Code for SPTO example: "PilotDataCommands_SPTO.txt"

```
# K. Irvine October 26, 2009
# KLMN Landbirds Protocol Development
# Importing Pilot Data and Extracting relevant values for power analysis
# Running an example for a power analysis based on pilot data from lava beds and SPTO (Spotted Towhee)
#######################################################################

#**************************
# STEP 1. IMPORT DATA
# The data file should be in a .csv file to import into R
# *****************************

Birds<-read.csv("PointCounts.csv",header=T)
names(Birds)

attach(Birds)

#*************
# STEP 2. SUBSET DATA TO DESIRED PARK, SPECIES
# subset to SPTO for Spotted Towhee data within Lava Beds example
#*************************

# 25 sites surveyed at Lava Beds
# 12 points each route

sites.SPTO<-SITE[PROJECT=="NPS_LABE"&SPECIES=="SPTO"& DISTANCE<=50]
points.SPTO<-POINT[PROJECT=="NPS_LABE"&SPECIES=="SPTO"& DISTANCE<=50]
#

# STEP 3.A CALCULATE RELATIVE ABUNDANCE FOR EACH ROUTE LAVA BEDS
table(sites.SPTO,points.SPTO)[1:25,] #calculating the number of detect for each point*site combination < 50 m
rel.abund.SPTO<-apply(table(sites.SPTO,points.SPTO)[1:25,],1,sum)/12 #taking the sum for a row divided by 12
points

# STEP 4. LOOK AT THE DATA TO ASSESS NORMALITY ON THE LOG-SCALE

hist(log(rel.abund.SPTO)) #check that looks like a "normal" distribution

# STEP 5. CALCULATE SITE-TO-SITE VARIANCE COMPONENTS AND OVERALL MEAN ACROSS SITES

sig2a<-var(log(rel.abund.SPTO)) #estimate of the site-to-site variability for log relative abundance
sig2a #0.4907695
mu<-mean(log(rel.abund.SPTO))
mu #-0.6430166

#
******************************************************************************************
******************#
```

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

```
#
                                          #
# STEP 6. EXAMPLE FOR ONE SPECIES: IN ORDER TO INVESTIGATE HOW MANY TRANSECTS AND
YEARS OF SAMPLING                  #
#  IS REQUIRED TO HAVE 80% POWER TO DETECT GIVEN LEVEL OF DECLINE/INCREASE OVER 20
YEARS
#
                                                 #
#*********************************************************************************************
********************#

#.10 = alpha level for trend test
# 500 is the number of iterations for each simulation
# c(20,25,30) is the number of sites
# 2 is the number of off years between sampling 0,3,6,9 for 10 years period of sampling
# c(5,10,20,30) is the number of years of sampling
# 2008 start year
# -50% decrease over 20 years
# mu is the grand mean
# sig2a is the variance component for random intercept
# sig2b is the year to year variability
# sig2t is the variance component for the random slopes
# sig2c is the error variance of site*year variance component

#We have no estimate of sig2b, sig2t, or sig2c. We can assume sig2b and sig2t are zero but we do need an estimate
of site*year variance
# from the Ashroot data anywhere between .05 to .01 might be reasonable guesses...

OUT<-POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-
.64,sig2a=.491,sig2b=0, sig2t=0,sig2c=.05)

#THE OUTPUT WILL BE A TABLE OF THE ESTIMATED POWER FOR A GIVEN NUMBER OF ROUTES
AND NUMBER OF YEARS OF SAMPLING [ROWS=NO.SITES, COLUMNS=NO.YEARS]
#WITH THE EXAMPLE ABOVE WE GOT

# 5  10  20  30
#20 0 0.814 0.972 0.980
#25 0 0.914 0.992 0.996
#30 0 0.956 0.994 1.000

OUT3<-POWER.RandomInterceptsModel(.10,200,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-
.64,sig2a=.491,sig2b=0, sig2t=0,sig2c=1)
OUT3

# 5  10  20  30
#20 0 0.315 0.74 0.845
#25 0 0.405 0.84 0.940
#30 0 0.515 0.93 0.985
```

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol.

Appendix 2. R-code for PSFL example

Code for PSFL example: "PilotDataCommands_PSFL.txt"

```
# K. Irvine October 26, 2009
# KLMN Landbirds Protocol Development
# Importing Pilot Data and Extracting relevant values for power analysis
# Running an example for a power analysis based on pilot data from redwood and Pacific-Slope Flycatcher
#######################################################################

#**************************
# STEP 1. IMPORT DATA
# The data file should be in a .csv file to import into R
# *****************************

Birds<-read.csv("PointCounts.csv",header=T)
names(Birds)

attach(Birds)

#*************
# STEP 2. SUBSET DATA TO DESIRED PARK, SPECIES
# subset to PSFL for Pacific slope flycatcher data within Redwoods example
#*************************

sites.PSFL<-SITE[PROJECT=="NPS_RNSP"&SPECIES=="PSFL"& DISTANCE<=50]
points.PSFL<-POINT[PROJECT=="NPS_RNSP"&SPECIES=="PSFL"& DISTANCE<=50]

# STEP 3.B CALCULATE RELATIVE ABUNDANCE FOR EACH ROUTE FOR REDWOODS

# 30 sites surveyed at Redwoods
# variable number of points each route

Routes<-read.csv("RouteandPoints.csv",header=T) #loading the route and points dataset
pts<-apply(table(Routes$ROUTE[Routes$PROJECT=="NPS_RNSP"],
Routes$STATION[Routes$PROJECT=="NPS_RNSP"])[27:56,],1,sum) #calculating the number of points surveyed each route
for all parks
pts

rel.abund.PSFL<-apply(table(sites.PSFL,points.PSFL)[27:56,],1,sum)/pts #taking the sum for a row divided by # points

# STEP 4. LOOK AT THE DATA TO ASSESS NORMALITY ON THE LOG-SCALE

hist(log(rel.abund.PSFL)) #check that looks like a "normal" distribution

# STEP 5. CALCULATE SITE-TO-SITE VARIANCE COMPONENTS AND OVERALL MEAN ACROSS SITES

sig2a<-var(log(rel.abund.PSFL)) #estimate of the site-to-site variability for log relative abundance
sig2a #.611
mu<-mean(log(rel.abund.PSFL))
mu #-.45

#
*********************************************************************************************************
*********#
#
                                                                #
# STEP 6. EXAMPLE FOR ONE SPECIES: IN ORDER TO INVESTIGATE HOW MANY TRANSECTS AND YEARS OF
SAMPLING                            #
```

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

IS REQUIRED TO HAVE 80% POWER TO DETECT GIVEN LEVEL OF DECLINE/INCREASE OVER 20 YEARS
#
                                        #
#***
**********#

#.10 = alpha level for trend test
500 is the number of iterations for each simulation
c(20,25,30) is the number of sites
2 is the number of off years between sampling 0,3,6,9 for 10 years period of sampling
c(5,10,20,30) is the number of years of sampling
2008 start year
-50% decrease over 20 years
mu is the grand mean
sig2a is the variance component for random intercept
sig2b is the year to year variability
sig2t is the variance component for the random slopes
sig2c is the error variance of site*year variance component

#We have no estimate of sig2b, sig2t, or sig2c. We can assume sig2b and sig2t are zero but we do need an estimate of site*year variance
from the Ashroot data anywhere between .05 to .01 might be reasonable guesses, but these seem low

OUT.PSFL1<-POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-.45,sig2a=.611,sig2b=0,
sig2t=0,sig2c=1)

#THE OUTPUT WILL BE A TABLE OF THE ESTIMATED POWER FOR A GIVEN NUMBER OF ROUTES AND
NUMBER OF YEARS OF SAMPLING [ROWS=NO.SITES, COLUMNS=NO.YEARS]
#WITH THE EXAMPLE ABOVE WE GOT

OUT.PSFL1

5 10 20 30
#20 0 0.132 0.402 0.598
#25 0 0.124 0.544 0.630
#30 0 0.148 0.584 0.744

# Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

MEMO 3: November 3, 2009

November 3, 2009
Prepared by K. Irvine, PhD

Dept of Math Sciences
Montana State University
Bozeman, MT 59715
irvine@math.montana.edu
(406) 994-5357

In this document, I show how to calculate the variance components for the power analysis for trend based on BBS survey routes located near the parks.

Again I assume that the routes are fixed.

Example 1. Spotted Towhee

The example used the pilot data provided by Sean Mohren: *BBS_SPTO_4_Routes_20091102.xls.* The R code is provided in *VarianceComponentEstimates.txt*

Step 1: In Excel

I made the following changes in excel:
1) Changed dashes to "NA"
2) Saved as a comma delimited file

Step 2: Imported into R

Once imported into R, I subset to only the relevant rows, for some reason it was reading in a much larger file.

Step 3: Plotted the pilot data
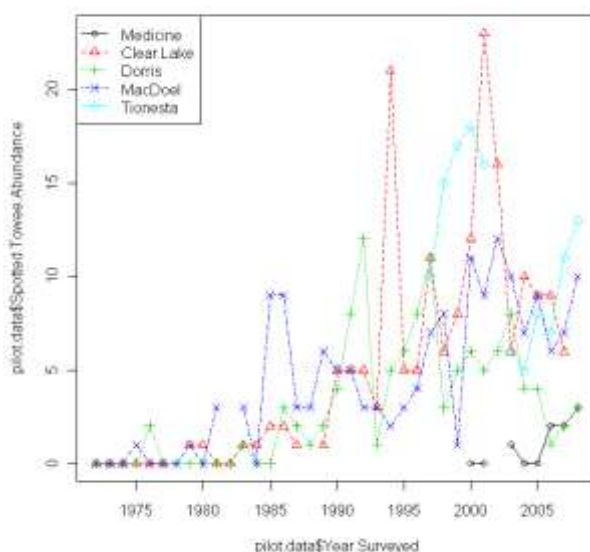
Figure 6. Plot of the abundance of Spotted Towhee along five Breeding Bird Survey routes in the proximity of LABE.

There is a lot of year-to-year variability and site-to-site variability, as expected.
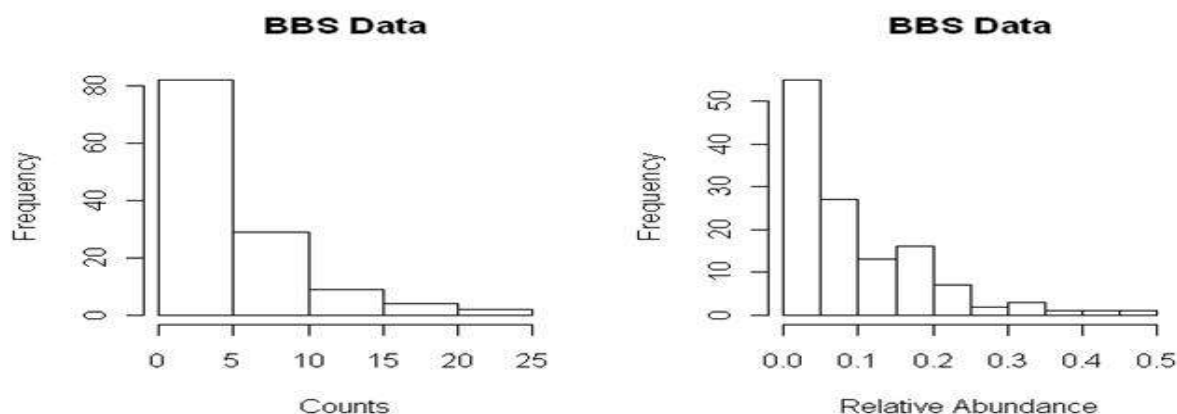
Figure 7. Histograms examining the year-to-year and site-to-site variability of the number and relative abundance of Spotted Towhee at five breeding bird survey routes.

Step 4: Calculating log transformed Relative Abundance

I divide by 50 stops for each count. Then I take the natural log transformation, ln(relative abundance+.001).

Step 5: Fit model to estimate variance components

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

```
> fit<- lmer(log.RE~Year.Surveyed +(1|Route.Surveyed)+(1|Year.Surveyed), data=pilot.data
)
> summary(fit)
Linear mixed model fit by REML
Formula: log.RE ~ Year.Surveyed + (1 | Route.Surveyed) + (1 | Year.Surveyed)
  Data: pilot.data
  AIC  BIC logLik deviance REMLdev
 438.7 452.9 -214.4   423  428.7
Random effects:
 Groups    Name    Variance Std.Dev.
 Year.Surveyed (Intercept) 0.3696  0.60794
 Route.Surveyed (Intercept) 2.2257  1.49188
 Residual          1.2084  1.09929
Number of obs: 126, groups: Year.Surveyed, 37; Route.Surveyed, 5

Fixed effects:
     Estimate Std. Error t value
(Intercept)  -313.82704  27.14663 -11.56
Year.Surveyed  0.15558  0.01362  11.42

Correlation of Fixed Effects:
    (Intr)
Year.Survyd -1.000
```

Just making sure the residuals appear relative normal…

Step 6: Estimating Power for based on assumed variance component estimates

In the power code, we would use
            #mu=-313
                    # sig2a is the variance component for random intercept
                    # sig2b is the year to year variability
                    # sig2t is the variance component for the random slopes
                    # sig2c is the error variance of site*year variance component

sig2a=2.22, sig2b=.37, and sig2c=1.099 based on this output.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).
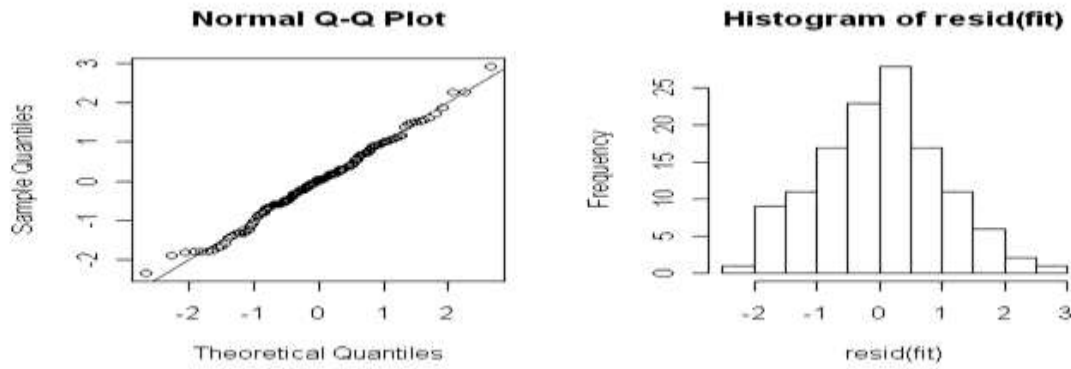

**Figure 8.** Examination of the normality of the residuals using Q-Q plots and histograms.

OUT<-POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-313,sig2a=2.22,sig2b=.37, sig2t=0,sig2c=1.1)

Sites	10	20	30	40
20	0.122	0.424	0.584	0.682
25	0.094	0.398	0.536	0.654
30	0.086	0.376	0.564	**0.654**

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

Power to Detect a -50% Trend after 20 years 2 years off between
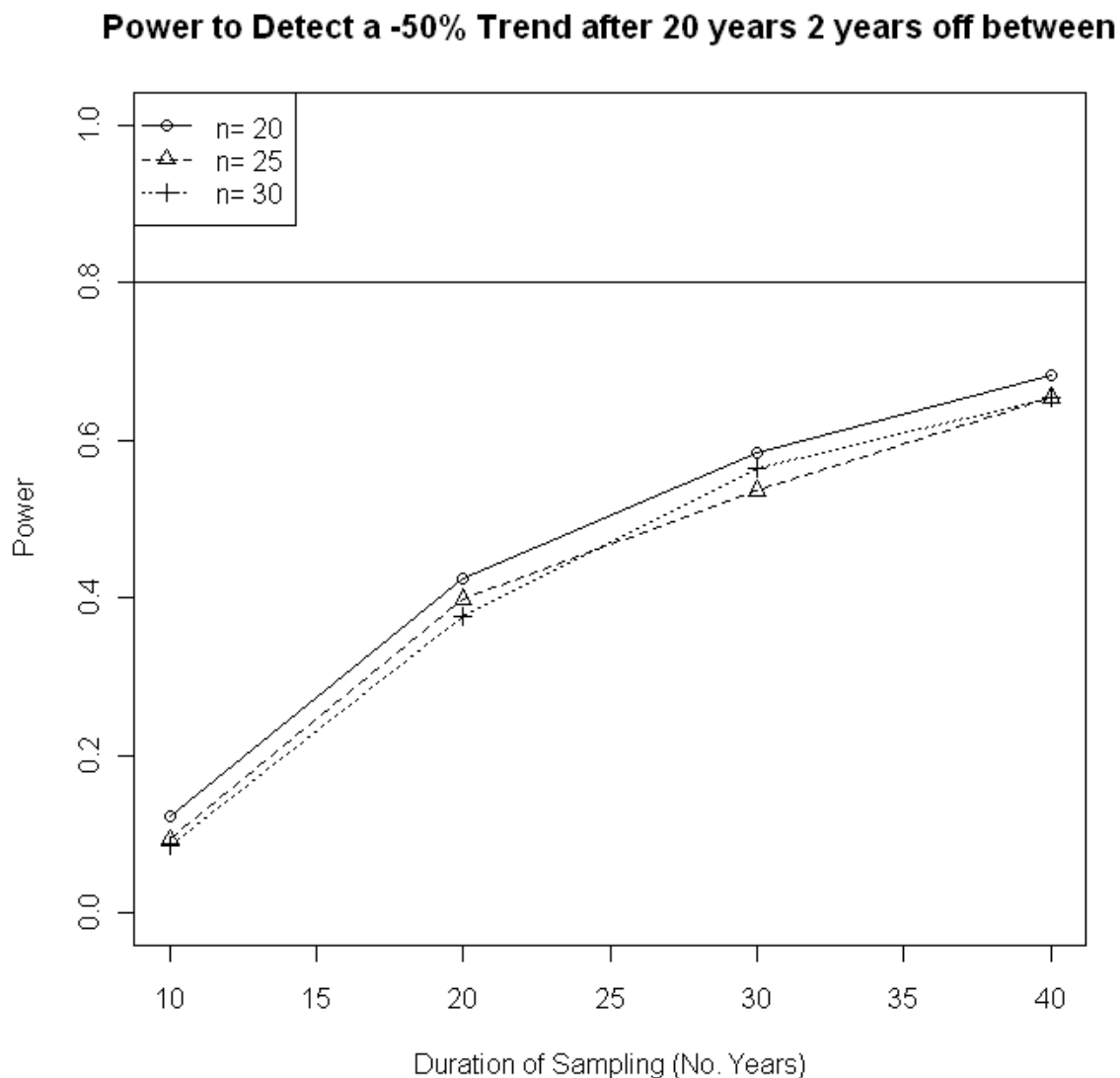
Figure 9. Power to detect trends in Spotted Towhee over a 30 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 0.05.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

Example 2. Brown Headed Cowbird

Plot of the BBS data:
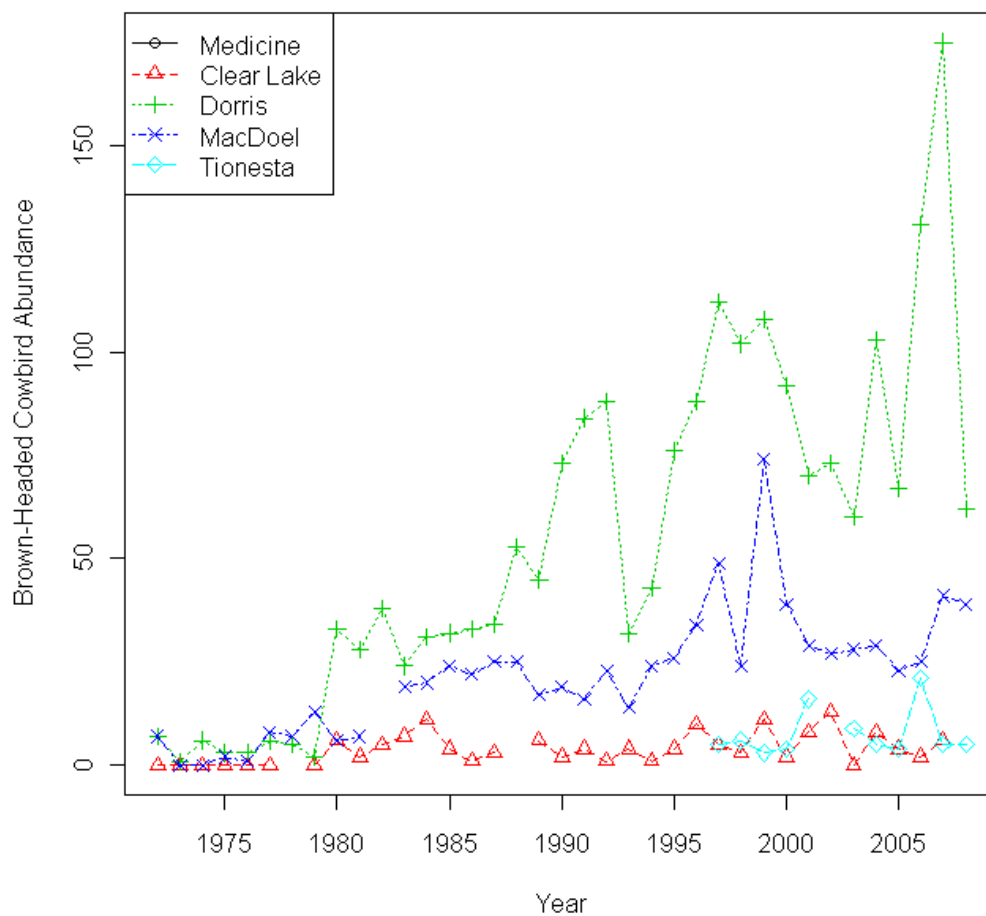
Figure 10. Plot of the abundance of Brown Headed Cowbird along five Breeding Bird Survey routes in the proximity of LABE.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).
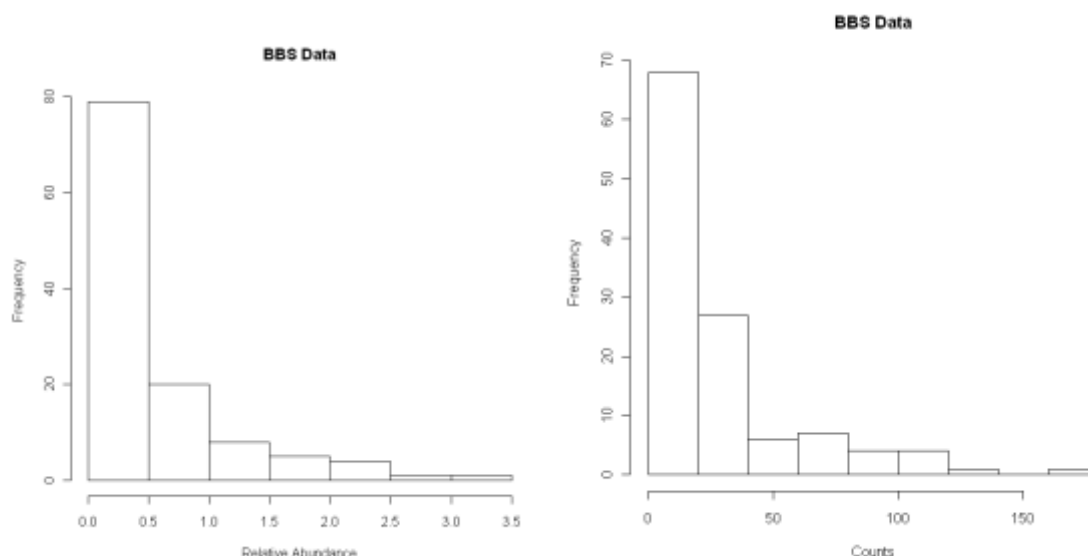
Figure 11. Histograms examining the year-to-year and site-to-site variability of the number and relative abundance of Brown Headed Cowbird at five breeding bird survey routes.

Results from fitting the mixed model for trend:

Linear mixed model fit by REML
Formula: log.RE ~ Year.Surveyed + (1 | Route.Surveyed) + (1 | Year.Surveyed)
 Data: pilot.data

Random effects:
 Groups   Name   Variance Std.Dev.
 Year.Surveyed (Intercept) **0.59220** 0.76955
 Route.Surveyed (Intercept) **2.12162** 1.45658
 Residual   **0.81467** 0.90259
Number of obs: 118, groups: Year.Surveyed, 37; Route.Surveyed, 4

Fixed effects:
      Estimate Std. Error t value
(Intercept) -206.04608 28.66948 -7.187
Year.Surveyed 0.10249 0.01439 7.120

Notice we would interpret the trend as an 11% annual increase (1.11=exp(0.10249)) in the median relative abundance of Brown Headed Cowbirds for these 4 routes surveyed.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).
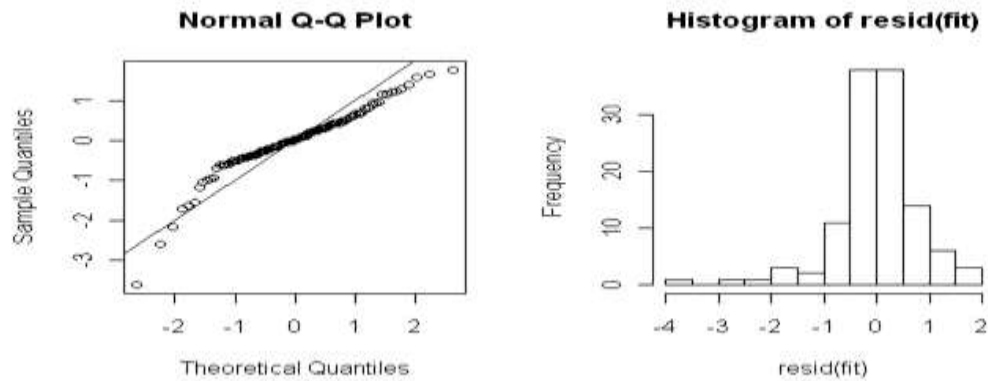

Figure 12. Examination of the normality of the residuals using Q-Q plots and histograms.

OUT<-POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(5,10,20,30),2008,-.5,mu=-206,sig2a=2.12,sig2b=.59, sig2t=0,sig2c=.82)

The residuals appear alright with the log-transformed plus .001.

OUT

Sites	5	10	20	30
20	0.004	0.118	0.438	0.582
25	0.000	0.100	0.442	0.548
30	0.000	0.112	0.432	**0.588**

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

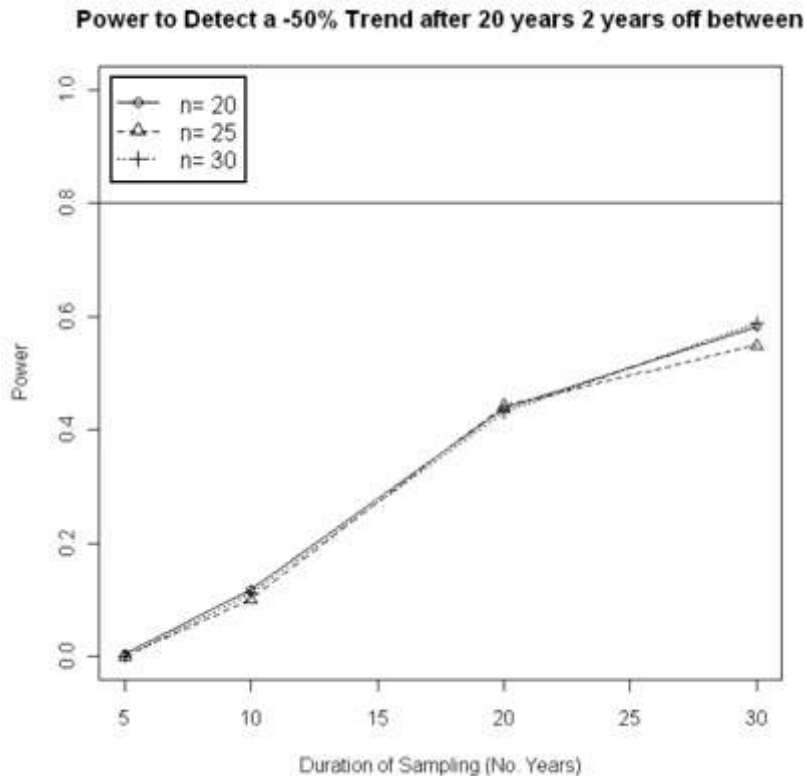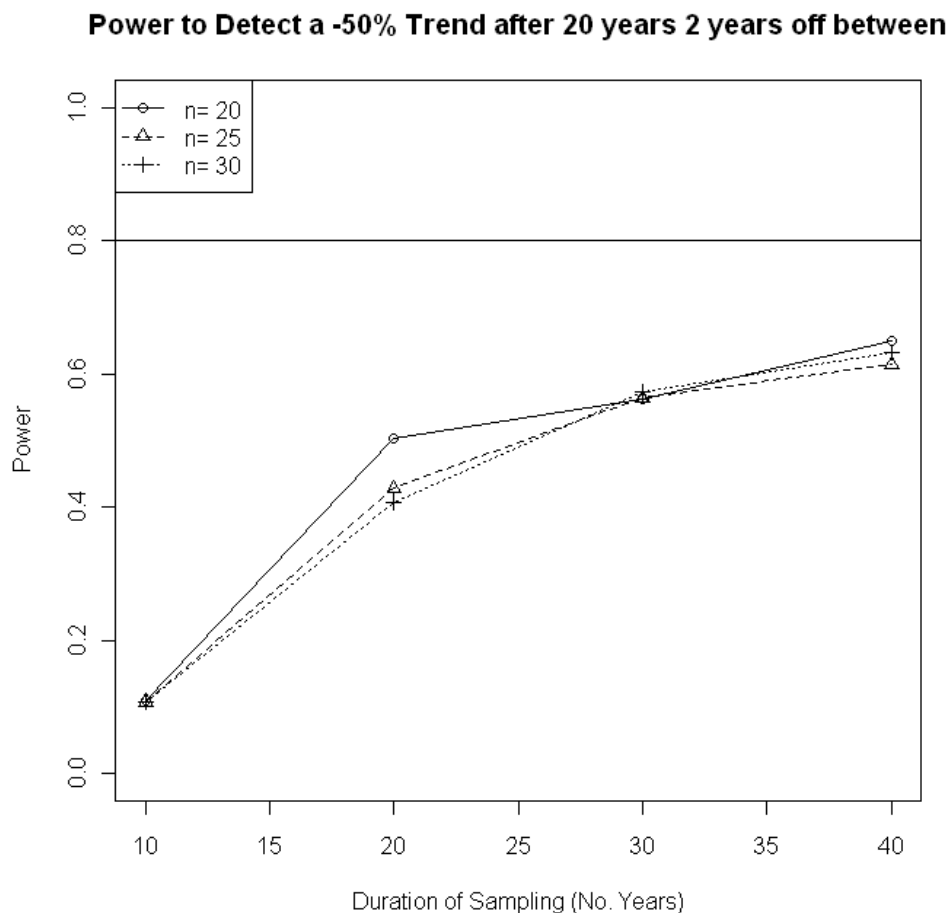**Power to Detect a -50% Trend after 20 years 2 years off between**

Figure 13. Power to detect trends in Brown Headed Cowbirds over a 30 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 0.05.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

Power to Detect a -50% Trend after 20 years 2 years off between

Figure 14. Power to detect trends in Brown Headed Cowbirds over a 40 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 0.05.

Example 3. Western Meadowlark
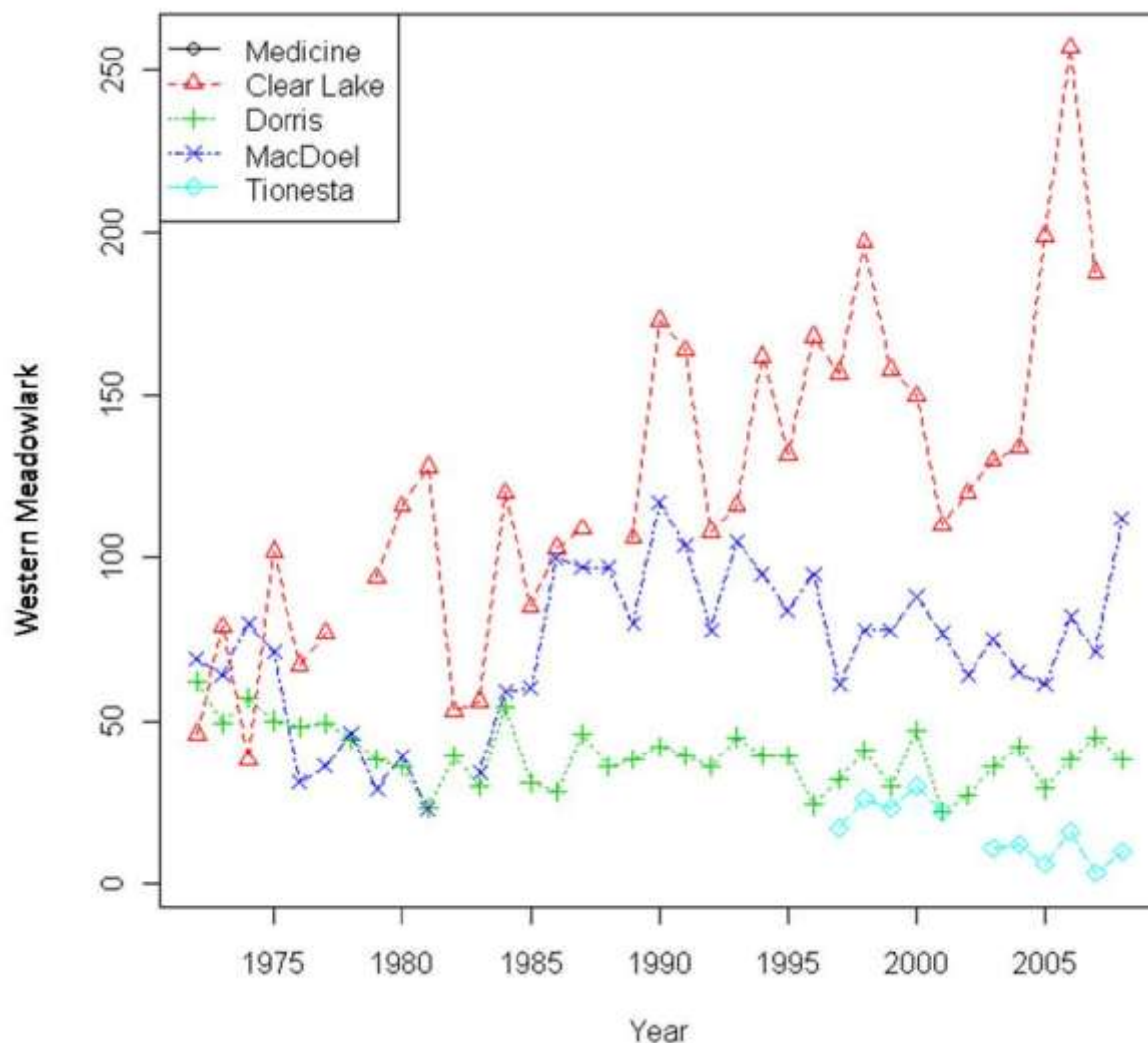
Plot of the BBS data:

Figure 15. Plot of the abundance of Western Meadowlarks along five Breeding Bird Survey routes in the proximity of LABE.

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).
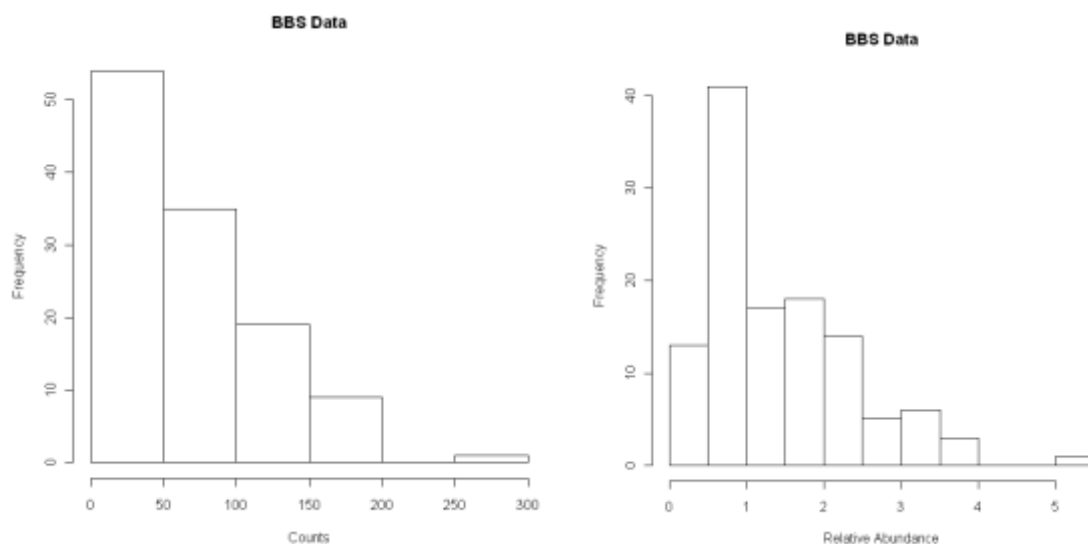

Figure 16. Histograms examining the year-to-year and site-to-site variability of the number and relative abundance of Western Meadowlarks at five breeding bird survey routes.
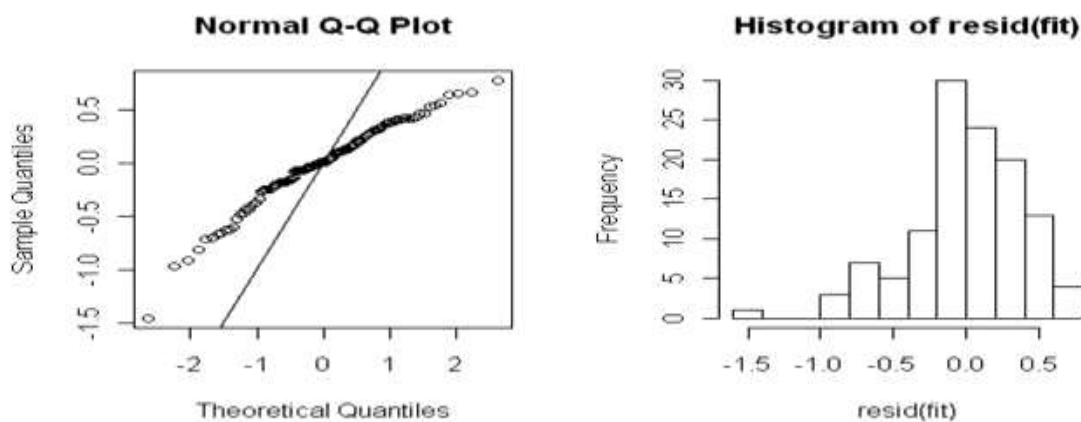.

Figure 17. Examination of the normality of the residuals using Q-Q plots and histograms.

Results from fitting the mixed model for trend:

The residuals appear not that great with the log-transformed plus .001.

Linear mixed model fit by REML
Formula: log.RE ~ Year.Surveyed + (1 | Route.Surveyed) + (1 | Year.Surveyed)
  Data: pilot.data
  AIC BIC logLik deviance REMLdev
 147.7 161.5 -68.83 128.5 137.7
Random effects:

Appendix J. Power Analysis Modeling of Selected Species for the KLMN Landbird Monitoring Protocol (continued).

```
 Groups    Name   Variance  Std.Dev.
 Year.Surveyed (Intercept) 1.0699e-15 3.2710e-08
 Route.Surveyed (Intercept) 9.0112e-01 9.4927e-01
 Residual        1.4894e-01 3.8593e-01
Number of obs: 118, groups: Year.Surveyed, 37; Route.Surveyed, 4

Fixed effects:
     Estimate Std. Error t value
(Intercept)  -21.330410  6.958447 -3.065
Year.Surveyed  0.010654  0.003483  3.059

OUT.WEME.BBS<-
POWER.RandomInterceptsModel(.10,500,c(20,25,30),2,c(10,20,30,40),2008,-.5,mu=-
21,sig2a=.901,sig2b=0, sig2t=0,sig2c=.15)
```

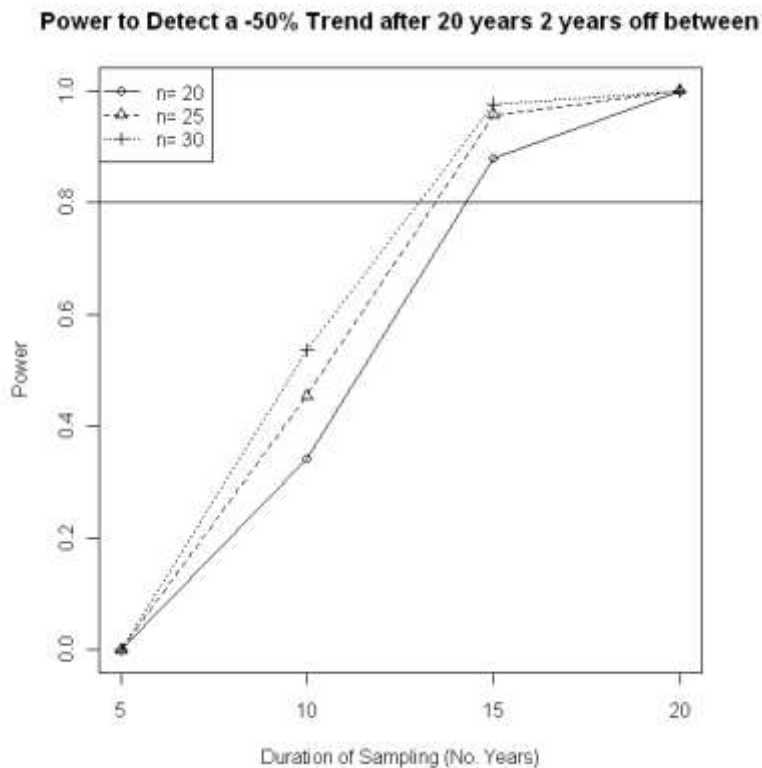| Sites	5	10	15	20
20	0	0.342	**0.878**	1
25	0	0.454	0.956	1
30	0	0.536	0.976	1

Figure 18. Power to detect trends in Western Meadowlarks over a 20 year period using 20, 25, and 30 routes at Lava Beds National Monument using an estimated variance component equal to 0.05.